EQUILIBRIUM PROBLEMS AND PROXIMAL ALGORITHMS IN HADAMARD SPACES

POOM KUMAM AND PARIN CHAIPUNYA*

KMUTT-Fixed Point Theory and Applications Research Group (KMUTT-FPTARG)
Theoretical and Computational Science Center (TaCS), Science Laboratory Building,
Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT),
126 Pracha-Uthit Road, Bang Mod, Thrung Khru, Bangkok 10140, Thailand.
Email: poom.kum@kmutt.ac.th

ABSTRACT. In this paper, we consider the equilibrium problems and also their regularized problems under the setting of Hadamard spaces. The solution to the regularized problem is represented in terms of resolvent operators. As a piece of essential machinery in the existence of an equilibrium, we first prove that the KKM principle is attained in general Hadamard spaces without assuming the compactness of the closed convex hull of a finite set. We construct the proximal algorithm based on this regularization and give convergence analysis adequately.

KEYWORDS: Equilibrium problems, Proximal algorithms, KKM principle, Hadamard space.
AMS Subject Classification: : 90C33, 65K15, 49J40, 49M30, 47H05.

1. INTRODUCTION

Equilibrium problems were originally studied in [8] as a unifying class of variational problems. Given a nonempty set K, and a bifunction $F : K \times K \to \mathbb{R}$. The equilibrium problem $EP(K, F)$ is formulated as follows:

Find a point $\bar{x} \in K$ such that $F(\bar{x}, y) \geq 0$ for every $y \in K$. $EP(K, F)$

By assigning different settings to F, we can include, *e.g.,* minimization, minimax inequalities, variational inequalities, and fixed point problems in the class of equilibrium problems. The set of all solutions of $EP(K, F)$ is denoted by $\mathcal{E}(K, F)$. Typical studies for $EP(K, F)$ are extensively carried out in Banach or Hilbert spaces, and recently in Hadamard manifolds.

Proximal algorithm is one of the most elementary method used in solving several classes of variational problems. The main idea of the method is to perturb (or

---

* *Corresponding author.* Parin Chaipunya .
Email address : parin.cha@mail.kmutt.ac.th.

regularize) the original problem into a simpler and more well-behaved problem, and solve for each steps the perturbed subproblems. This method was originally proposed under the setting of Hilbert spaces by Martinet [24], and was progressively developed by Rockafellar [30].

Proximal algorithms in Hadamard spaces were recently investigated. In particular, it was introduced for minimizing convex functionals by Bačák in [5], and were extended for solving variational inequalities in [17] and for solving zeros of maximal monotone operators in [18, 10].

Equilibrium problems in Hadamard manifold were studied by [11] and were later extended to equilibrium problems for bifunctions defined on proximal pairs in [9]. Both of the results rely on different variants of the KKM lemma (consult [22] for the original version), but the latter is strongly based on Brouwer's fixed point theorem in Hadamard manifolds (see [27]). In fact, the Brouwer's theorem is proved in Hadamard spaces in [29, 28] and the KKM principle was subsequently proved in Hadamard space with the convex hull finite property (CHFP), *i.e.,* every polytope is assumed to be compact (see also the convex hull property in [15]).

Based on this KKM principle and under CHFP assumption, Kimura and Kishi [19] studied the equilibrium problem and showed the well-definedness of the resolvent associated to a bifunction. In particular, they also proved that this resolvent is firmly nonspreading and has fixed point set identical to the equilibrium points. Finally, they apply a convergence theorem from [20] to solve for an equilibrium.

In this paper, we show that the KKM principle can be extended to any Hadamard spaces without further assumptions. Using the KKM principle, we show the existence of an equilibrium of a bifunction under standard continuity, convexity, and compactness/coercivity assumptions. Then, we deduce several fundamental properties of the resolvent operator introduced by Kimura and Kishi [19]. We also deliver some comparisons with resolvents of convex functionals and of monotone vector fields. We finally define the proximal algorithm by iterating resolvent operators and provide adequate convergence analysis of the algorithm. Apart from dropping the CHFP assumption, the contents presented in this paper provide a continuation from the works of Niculescu and Rovenţa [29, 28] in the study of KKM principle, of Kimura and Kishi [19] in the study of bifunctions, their resolvents, and proximal algorithms, and of the authors [10] in the relationships between bifunctions and vector fields as well as their resolvents and convergence results.

The organization of this paper is as follows. The next section collects useful basic knowledges used in the rest of this paper. We also give in this section several auxiliary results that will be exploited to validate subsequent results. In particular, properties concerning the product $\langle \cdot, \cdot \rangle$ and the weaker convergence notions are explained. Section 3 is devoted to the discussion on the KKM theory on Hadamard spaces. This contains the key tool for proving existence theorems, which will be found in Section 4. We also deduce here the dual problem in the sense of Minty and give relationships with the primal priblem. In Section 5, we introduce the resolvent operator corresponds to a bifunction. Properties of the resolvents are thoroughly deduced, especially with the nonexpansivity of the operator. This leads to the study of Section 7, which contains the construction of the proximal algorithm and also its convergence analysis.

2. Preliminaries and Auxiliaries

We divide this section into several parts, describing each topics in brief details. This includes some basic definitions and natations up to the technical results that will be used in our main results in the next sections.

2.1. **Hadamard spaces.** A uniquely geodesic metric space (X, ρ) is a CAT(0) space if each geodesic triangle in X is at least as thin as its comparison triangle in Euclidean plane. A complete CAT(0) space is then called *Hadamard space*. The following characterization of a CAT(0) space is useful.

Proposition 2.1. *For a geodesic metric space (X, ρ), the following conditions are all equivalent:*

(i) X is CAT(0).
(ii) ([13]) For any $x, u, v \in X$ and $\lambda \in [0, 1]$, the following inequality holds:

$$\rho^2(x, \gamma_{u,v}(\lambda)) \le (1 - \lambda)\rho^2(x, u) + \lambda\rho^2(x, v) - \lambda(1 - \lambda)\rho^2(u, v). \qquad \text{(CN)}$$

(iii) ([7]) For any $x, y, u, v \in X$, the following inequality holds:

$$\rho^2(x, v) + \rho^2(y, u) \le \rho^2(x, u) + \rho^2(y, v) + 2\rho(x, y)\rho(u, v). \qquad (2.1)$$

Definition 2.2. A convex set $K \subset X$ is said to be *flat* if the (CN) inequality holds as an equality for each $x, u, v \in K$.

Let (X, ρ) be an Hadamard space. For each $x, y \in X$, we write $\gamma_{x,y} : [0, 1] \to X$ to denote the normalized geodesic joining x and y, i.e., $\gamma_{x,y}(0) = x$, $\gamma_{x,y}(1) = y$, and $\rho(\gamma_{x,y}(t), \gamma_{x,y}(t')) = \rho(x, y) |t - t'|$ for all $t, t' \in [0, 1]$. We also adopt the following notations: $[\![x, y]\!] := \{\gamma_{x,y}(t); t \in [0, 1]\}$ and $(1 - t)x \oplus ty := \gamma_{x,y}(t)$ for $t \in [0, 1]$. A subset $K \subset X$ is said to be *convex* if $[\![x, y]\!] \subset K$ for any $x, y \in K$, and for a set $E \subset X$ we write co(E) to denote the smallest convex set that contains E. Certainly, we call co(E) the *convex hull* of E. A function $h : K \to \mathbb{R}$, with K being convex, is called *convex* (resp., *quasi-convex*) if $h \circ \gamma_{x,y} : [0, 1] \to \mathbb{R}$ is convex (resp., *quasi-convex*) for any $x, y \in K$. Moreover, h is called *concave* (resp., *quasi-concave*) if $-h$ is convex (resp., *quasi-convex*).

If (X, d) and (X', d') are two Hadamard space, then the product $X \times X'$ is also an Hadamard space with the metric given by

$$\rho\left((x, x'), (y, y')\right) := \left(d^2(x, y) + d'^2(x', y')\right)^{\frac{1}{2}}, \quad \forall (x, x'), (y, y') \in X \times X'.$$

Unless otherwise stated, always assume throughout this paper that (X, ρ) is an Hadamard space and $K \subset X$ is nonempty, closed, and convex. Any product space is also to be understood to be an Hadamard space in the sense described above.

2.2. **Dual space and Tangent spaces.** The concept of a dual space of X was introduced in [2]. Here, we recall such a construction in brief details. Let us write $\overrightarrow{xy} := (x, y) \in X^2$. The *quasilinearization* (see [6, 7]) on X^2 is the product $\langle \cdot, \cdot \rangle : X^2 \times X^2 \to \mathbb{R}$ defined by

$$\langle \overrightarrow{uv}, \overrightarrow{xy} \rangle := \frac{1}{2}[\rho^2(u, y) + \rho^2(v, x) - \rho^2(u, x) - \rho^2(v, y)], \quad \forall \overrightarrow{uv}, \overrightarrow{xy} \in X^2.$$

For each $t \in \mathbb{R}$ and $\overrightarrow{uv} \in X^2$, we define the Lipschitzian function $\Theta(t; \overrightarrow{uv}) : X \to \mathbb{R}$ by

$$\Theta(t; \overrightarrow{uv})(x) := t\langle \overrightarrow{uv}, \overrightarrow{ux} \rangle, \quad \forall x \in X.$$

Recall that for a Lipschitzian function $h : X \to \mathbb{R}$, the Lipschitzian constant of h is the quantity $L(h) := \inf\{\frac{|h(x)-h(y)|}{\rho(x,y)} \,|\, x, y \in X,\, x \neq y\}$. Applying the Lipschizian constant to the differences of Θ's, we can define a pseudometric $\tilde{D}$ on $\mathbb{R} \times X^2$ by:

$$\tilde{D}\left((s, \overrightarrow{uv}), (t, \overrightarrow{xy})\right) := L(\Theta(s; \overrightarrow{uv}) - \Theta(t; \overrightarrow{xy})), \quad \forall (s, \overrightarrow{uv}), (t, \overrightarrow{xy}) \in \mathbb{R} \times X^2,$$

which naturally gives rise to the following equivalence relation

$$(s, \overrightarrow{uv}) \sim (t, \overrightarrow{xy}) \iff \tilde{D}\left((s, \overrightarrow{uv}), (t, \overrightarrow{xy})\right) = 0, \quad \forall (s, \overrightarrow{uv}), (t, \overrightarrow{xy}) \in \mathbb{R} \times X^2.$$

The quotient space $X^* := \mathbb{R} \times X^2 / \sim$ with a metric D given by

$$D([(s, \overrightarrow{uv})]_\sim, [(t, \overrightarrow{xy})]_\sim) = \tilde{D}\left((s, \overrightarrow{uv}), (t, \overrightarrow{xy})\right), \quad \forall [(s, \overrightarrow{uv})]_\sim, [(t, \overrightarrow{xy})]_\sim \in X^*,$$

is called the *dual space* of X. For simplicity, we adopt the notation $s\overrightarrow{uv} := [(s, \overrightarrow{uv})]_\sim$. Moreover, we write $\mathbf{0} := s\overrightarrow{uu} = 0\overrightarrow{uv}$ for $s \in \mathbb{R}$ and $u, v \in X$.

In this paper, we restrict to particular subspaces of X^*. For any given $p \in X$, the *tangent space* of X at p is given by $T_p X := \{s\overrightarrow{py} \,|\, s \geq 0,\, y \in X\}$.

The following properties of the quasilinearization are fundamental.

Lemma 2.3. *For every $u, v, z \in X$, the following properties hold:*

(i) $\langle \overrightarrow{uv}, \overrightarrow{uz} \rangle + \langle \overrightarrow{vu}, \overrightarrow{vz} \rangle = \rho^2(u,v) \geq 0$.

(ii) *If $\lambda \in [0,1]$ and $x = \gamma_{z,u}(\lambda)$, it holds that $\lambda \langle \overrightarrow{zu}, \overrightarrow{zv} \rangle \leq \langle \overrightarrow{zx}, \overrightarrow{zv} \rangle$. In addition, if a convex set $K \subset X$ is flat, then the above inequality becomes equality for $u, v, z \in K$.*

(iii) *If a convex set $K \subset X$ is flat and $z, v \in K$, then $u \in K \mapsto \langle \overrightarrow{zu}, \overrightarrow{zv} \rangle$ is affine on K, i.e., it is both convex and concave on K.*

Proof. The proof for (i) is trivial. For (ii), the (CN) inequality gives

$$\begin{aligned}
2\langle \overrightarrow{zx}, \overrightarrow{zv} \rangle &= \rho^2(z,x) + \rho^2(z,v) - \rho^2(x,v) \\
&= \lambda^2 \rho^2(z,u) + \rho^2(z,v) - \rho^2(\gamma_{z,u}(\lambda), v) \\
&\geq \lambda^2 \rho^2(z,u) + \rho^2(z,v) \\
&\quad - \left[(1-\lambda)\rho^2(z,v) + \lambda\rho^2(u,v) - \lambda(1-\lambda)\rho^2(z,u)\right] \\
&= \lambda\rho^2(z,u) + \lambda\rho^2(z,v) - \lambda\rho^2(u,v) \\
&= 2\lambda\langle \overrightarrow{zu}, \overrightarrow{zv} \rangle,
\end{aligned}$$

which proves the first assertion. Now, if a convex set $K \subset X$ is flat, then the (CN) inequatily becomes equality for $u, v, z \in K$, and the desired result is straightforward from the above proofline.

Next, let us show (iii). Let $p, q \in K$ and $\lambda \in [0,1]$. It follows that

$$\begin{aligned}
2\langle \overrightarrow{z\gamma_{p,q}(\lambda)}, \overrightarrow{zv} \rangle &= \rho^2(z, \gamma_{p,q}(\lambda)) + \rho^2(z,v) - \rho^2(\gamma_{p,q}(\lambda), v) \\
&= \left[(1-\lambda)\rho^2(z,p) + \rho^2(z,q) - \lambda(1-\lambda)\rho^2(p,q)\right] + \rho^2(z,v) \\
&\quad - \left[(1-\lambda)\rho^2(v,p) + \lambda\rho^2(v,q) - \lambda(1-\lambda)\rho^2(p,q)\right] \\
&= (1-\lambda)\left[\rho^2(z,p) + \rho^2(z,v) - \rho^2(p,v)\right] \\
&\quad + \lambda\left[\rho^2(z,q) + \rho^2(z,v) - \rho^2(q,v)\right] \\
&= 2 \cdot \left[(1-\lambda)\langle \overrightarrow{zp}, \overrightarrow{zv} \rangle + \lambda\langle \overrightarrow{zq}, \overrightarrow{zv} \rangle\right].
\end{aligned}$$

The proof is thus completed.   ∎

2.3. **Modes of convergence.** Convergence in the metric topology is known to be irrelevant in some situations, especially in the study of numerical algorithms in infinite dimensional spaces. In this subsection, we recall two alternative modes of convergence for bounded sequences, namely the Δ- and w-convergences. Both of the concepts are identical to weak convergence in Hilbert spaces.

Let us start with the Δ-convergence. Suppose that $(x^k) \subset X$ be a bounded sequence, and define a function $r(\cdot; (x^k)) : X \to [0, \infty)$ by

$$r(x; (x^k)) := \limsup_{k \to \infty} \rho(x, x^k), \quad \forall x \in X.$$

The minimizer of this function is known to exists and is unique (see [12]). Following [21] (see also [23]), a bounded sequence (x^k) is said to be Δ-*convergent* to a point $\bar{x} \in X$ if $\bar{x} = \arg\min_{x \in X} r(x; (u^k))$ for any subsequence $(u^k) \subset (x^k)$. In this case, $\bar{x}$ is called the Δ-limit of (x^k). Recall that a bounded sequence is Δ-convergent to at most one point.

The topology that generates the Δ-convergence is unknown in general. However, we still adopt the topological-like notions such as Δ-accumulation points, Δ-closed sets, or Δ-continuity. For instance, a point $u \in X$ is a Δ-*accumulation point* of the sequence $(x^k) \subset X$ if it contains a subsequence that is Δ-convergent to u. A set $K \subset X$ is called Δ-*closed* if each Δ-convergent sequence in K has its Δ-limit in K. A function $f : X \to \mathbb{R}$ is called Δ-*upper semicontinuous* (briefly Δ-*usc*) if its epigraph is Δ-closed in $X \times \mathbb{R}$.

The following proposition gives two most important properties regarding the Δ-convergence that are required in our main theorems.

Proposition 2.4. *Suppose that $(x^k) \subset X$ is bounded. Then, the following properties hold:*

    (i) ([23]) (x^k) has a Δ-convergent subsequence .

    (ii) ([1]) (x^k) is Δ-convergent to $\bar{x} \in X$ if and only if $\limsup_k \langle S, \overrightarrow{\bar{x}x^k} \rangle \leq 0$ for any $S \in T_{\bar{x}} X$.

Next, let us turn to the notion of w-convergence as introduced in [1]. A bounded sequence $(x^k) \subset X$ is said to be w-*convergence* to a point $\bar{x} \in X$ if $\lim_k \langle S, \overrightarrow{\bar{x}x^k} \rangle = 0$ for any $S \in T_{\bar{x}} X$. With Proposition 2.4, we can see immediately that w-convergence implies Δ-convergence. As was noted in [1, 3], a bounded sequence does not necessarily have a w-convergent subsequence. This motivates the definition of reflexivity in Hadamard spaces, i.e., a convex set $K \subset X$ is said to be *reflexive* if each bounded sequence in K contains a w-convergent subsequence. However, it turns out that the reflexivity of K implies the equivalence between the Δ- and weak convergences. Still, the reflexivity can be useful in obtaining sharper estimates in some situations (see e.g. the proof of Lemma 7.7).

Proposition 2.5 ([4]). *A convex set $K \subset X$ is closed if and only if it is Δ-closed.*

Proposition 2.6. *Suppose that a convex set $K \subset X$ is reflexive and a bounded sequence $(x^k) \subset K$ is Δ-convergent. Then, it is w-convergent.*

Proof. Suppose that (x^k) is Δ-convergent to $\bar{x} \in X$. Let us assume to the contrary that (x^k) is not weakly convergent. Equivalently, there must exists $\gamma_0 \in T_{\bar{x}} X$ such that

$$\liminf_{k \to \infty} \langle \overrightarrow{\bar{x}x^k}, \gamma_0 \rangle < \limsup_{k \to \infty} \langle \overrightarrow{\bar{x}x^k}, \gamma_0 \rangle \leq 0. \tag{2.2}$$

Suppose that $(x^{k_j}) \subset (x^k)$ is a subsequence such that

$$\lim_{j \longrightarrow \infty} \langle \overrightarrow{\bar{x}x^{k_j}}, \gamma_0 \rangle = \liminf_{k \longrightarrow \infty} \langle \overrightarrow{\bar{x}x^k}, \gamma_0 \rangle.$$

Take into account the inequality (2.2), we get $\lim_j \langle \overrightarrow{\bar{x}x^{k_j}}, \gamma_0 \rangle < 0$, which prevents (x^k) from having a weakly convergent subsequence. Since (x^k) is bounded in K, this violates the reflexivity of K. Therefore, (x^k) must be weakly convergent to $\bar{x}$.
∎

2.3.1. *Fejér convergence.* Lastly, the notion of Fejér convergence is essential and will play a central role in the proof of our main convergence theorems. This notion encapsulates the improvement at each iteration of some approximate sequence towards a solution set.

Definition 2.7. A sequence $(x^k) \subset X$ is said to be *Fejér convergent* with respect to a nonempty set $V \subset X$ if for each $x \in V$, we have $\rho(x^{k+1}, x) \le \rho(x^k, x)$ for all large $k \in \mathbb{N}$.

Proposition 2.8 ([10]). *Suppose that $(x^k) \subset X$ is Fejér convergent to a nonempty set $V \subset X$. Then, the following are true:*
  (i) *(x^k) is bounded.*
  (ii) *$(\rho(x, x^k))$ converges for any $x \in V$.*
  (iii) *If every Δ-accumulation point lies within V, then (x^k) is Δ-convergent to an element in V.*

3. The KKM Principle

The KKM principle was initiated in [22], and was successfully extended into topological vector spaces by Fan in [14]. Further extensions into Hadamard manifold was discussed in [11, 9]. The results can be generalized instantly also into Hadamard spaces with fixed point property for continuous mappings defined on a convex hull of finite points [28]. Here we prove the KKM principle without using such condition on the space. First, let us recall the original statement of [22] and another additional result.

Lemma 3.1 ([22]). *Let $C_1, \ldots, C_m$ be closed subsets of the standard $(m-1)$-simplex σ. If $\mathrm{co}(\{x_i \ ; \ i \in I\}) \subset \bigcup_{i \in I} C_i$ for each $I \subset \{1, \ldots, m\}$, then the intersection $\bigcap_{j=1}^{m} C_j$ is nonempty.*

Lemma 3.2. *Suppose that (x^k) is a sequence in X and define the following sequence of sets by induction:*

$$\begin{cases} D_1 := \{x_1\}, \\ D_j := \{z \in [\![x^j, y]\!] \ ; \ y \in D_{j-1}\}, \quad \text{for } j = 2, 3, \ldots \end{cases} \tag{3.1}$$

Then, D_j is compact for all $j \in \mathbb{N}$.

Proof. We shall prove the statement by using mathematical induction. It is obvious that D_1 is compact. Now, we enter the inductive step by assuming D_j is compact and show that D_{j+1} must be compact.

Suppose that (u^k) is an arbitrary sequence in D_{j+1}. By definition, there correspond sequences (v^k) in D_j and (t_k) in $[0, 1]$ such that

$$u^k = \gamma_{x^{j+1}, v^k}(t_k), \quad \forall k \in \mathbb{N}.$$

Since both D_j and $[0,1]$ are compact, we may find convergent subsequences (v^{k_i}) of (v^k) and (t_{k_i}) of (t_k) with limits $\bar{v} \in D_j$ and $\bar{t} \in [0,1]$, respectively. Set $\bar{u} := \gamma_{x^{j+1}, \bar{v}}(\bar{t})$, we now show that (u^{k_i}) is in fact convergent to $\bar{u}$.

Observe that

$$
\begin{aligned}
\rho(u^{k_i}, \bar{u}) &= \rho(\gamma_{x^{j+1}, v^{k_i}}(t_{k_i}), \gamma_{x^{j+1}, \bar{v}}(\bar{t})) \\
&\leq \rho(\gamma_{x^{j+1}, v^{k_i}}(t_{k_i}), \gamma_{x^{j+1}, \bar{v}}(t_{k_i})) + \rho(\gamma_{x^{j+1}, \bar{v}}(t_{k_i}), \gamma_{x^{j+1}, \bar{v}}(\bar{t})) \\
&\leq t_{k_i} \rho(v^{k_i}, \bar{v}) + |t_{k_i} - \bar{t}| \, \rho(x^{j+1}, \bar{v}).
\end{aligned}
$$

Passing $i \longrightarrow \infty$, we obtain from the above inequalities that $u^{k_i} \longrightarrow \bar{u}$. This guaratees the compactness of D_{j+1}, and the desired conclusion is thus proved. ∎

Theorem 3.3 (The KKM Principle). *Let $K \subset X$ be a closed convex set, $G : K \rightrightarrows K$ be a set-valued mapping with closed values. Suppose that for any finite subset $D := \{x_1, \ldots, x_m\} \subset K$, it holds the following inclusion:*

$$
\mathrm{co}(D) \subset G(D). \tag{3.2}
$$

Then, the family $\{G(x)\}_{x \in K}$ has the finite intersection property. Moreover, if $G(x_0)$ is compact for some $x_0 \in K$, then $\bigcap_{x \in K} G(x) \neq \varnothing$.

Proof. Let $D := \{x_1, \ldots, x_m\}$ be an arbitrary finite subset of K. For each $j = 1, 2, \ldots, m$, define D_j by using (3.1). Also, set $D^* := \bigcup_{j=1}^m D_j$.

Let $\sigma := \langle e_1, \ldots, e_m \rangle$ be the standard $(m-1)$-simplex in $\mathbb{R}^m$. Suppose that $\lambda_2 \in \langle e_1, e_2 \rangle$, then we can represent λ_2 with a scalar $s_1 \in [0,1]$ such that $\lambda = (1-s_1)e_2 + s_1 e_1$. Now, if $\lambda_3 \in \langle e_1, e_2, e_3 \rangle$, then $\lambda_3 = (1-s_2)e_3 + s_2 \lambda_2$ for some $\lambda_2 \in \langle e_1, e_2 \rangle$. By the earlier fact, λ_2 is represented by some $s_1 \in [0,1]$. Hence, λ_3 can be represented by a 2-dimensional vector $[s_1 \; s_2]^\top \in [0,1]^2$. Likewise for $j = 3, \ldots, m$, we can see that any $\lambda_j \in \langle e_1, \ldots, e_j \rangle$ is represented by a $(j-1)$-dimensional vector $[s_1 \; \ldots \; s_{j-1}]^\top \in [0,1]^{j-1}$. We shall adopt the notation $\lambda_j \equiv [s_1 \; \ldots \; s_{j-1}]^\top$ for the above representation.

Define a mapping $T : \sigma \to D^*$ by induction as follows: if $\lambda_2 \equiv s_1 \in \langle e_1, e_2 \rangle$, let $T(\lambda_2) := \gamma_{x_2, x_1}(s_1)$. For $j = 2, 3, \ldots, m$, if $\lambda_j \equiv [s_1 \; \ldots \; s_{j-1}]^\top \in \langle e_1, \ldots, e_j \rangle \setminus \langle e_1, \ldots, e_{j-1} \rangle$, then define $T(\lambda_j) := \gamma_{x_j, T(\lambda_{j-1})}(s_{j-1})$, where $\lambda_{j-1} \equiv [s_1 \; \ldots \; s_{j-2}]^\top \in \langle e_1, \ldots, e_{j-1} \rangle$.

We now show that T is continuous. Let $\lambda_m, \mu_m \in \sigma$, and $\lambda \equiv [s_1 \; \ldots \; s_{m-1}]^\top$ and $\mu \equiv [t_1 \; \ldots \; t_{m-1}]^\top$ respectively. For simplicity, let $\lambda_j \equiv [s_1 \; \ldots \; s_{j-1}]^\top$ and $\mu_j \equiv [t_1 \; \ldots \; t_{j-1}]^\top$, for $j = 1, \ldots, m$. Indeed, we have

$$
\begin{aligned}
\rho(T(\lambda_m), T(\mu_m)) &= \rho(\gamma_{x_m, T(\lambda_{m-1})}(s_{m-1}), \gamma_{x_m, T(\mu_{m-1})}(t_{m-1})) \\
&\leq \rho(\gamma_{x_m, T(\lambda_{m-1})}(s_{m-1}), \gamma_{x_m, T(\lambda_{m-1})}(t_{m-1})) \\
&\quad + \rho(\gamma_{x_m, T(\lambda_{m-1})}(t_{m-1}), \gamma_{x_m, T(\mu_{m-1})}(t_{m-1})) \\
&\leq |s_{m-1} - t_{m-1}| \operatorname{diam}(D^*) + \rho(T(\lambda_{m-2}), T(\mu_{m-2})) \\
&\;\;\vdots \\
&\leq \sum_{i=1}^{m-1} |s_i - t_i| \operatorname{diam}(D^*).
\end{aligned}
$$

This is sufficient to guarantee the continuity of T.

For each $j = 1, \ldots, m$, define $E_j := T^{-1}(D^* \cap G_i(x_j)) \subset \sigma$. By the continuity of T and Lemma 3.1, we may see that E_j's are closed sets. Suppose that $I \subset \{1, \ldots, m\}$,

$\lambda \in \text{co}(\{e_i \ ; \ i \in I\})$ and $\lambda \equiv [s_1 \ \ldots \ s_{m-1}]^\top$. Then, $s_j = 1$ if $j \notin I$. By the definition of T, we have

$$T(\lambda) \in \text{co}(\{x_i \ ; \ i \in I\}) \subset \bigcup_{i \in I} G(x_i).$$

It follows that $T(\lambda) \in D^* \cap G(x_i)$ for some $i \in I$. In other words, we have $\lambda \in E_i$ and therefore $\text{co}(\{e_i \ ; \ i \in I\}) \subset \bigcup_{i \in I} E_i$ for any $I \subset \{1, \ldots, m\}$. By applying the original KKM covering lemma, we get the existence of $\lambda^* \in \bigcap_{j=1}^m E_j$. Hence, we get $T(\lambda^*) \in \bigcap_{j=1}^m G(x_j)$. In fact, we have proved that the family of closed sets $\{G(x)\}_{x \in X}$ has the finite intersection property. The nonemptiness of the intersection $\bigcap_{x \in X} G(x) = \bigcap_{x \in X}(G(x) \cap G(x'))$ follows from the fact that $\{G(x) \cap G(x')\}_{x \in X}$ is a family of closed subsets with finite intersection property in a compact subspace $G(x')$. ∎

4. Solutions of Equilibrium Problems

We mainly discuss in this section two topics. First, we deduce the solvability of the problem $EP(K, F)$ under standard (semi)continuity, convexity, and compactness/coercivity assumptions. Secondly, we introduce the dual problem to $EP(K, F)$ in the sense of Minty and show the relationship between the primal and dual problems, in terms of their solutions. Note that the knowledge of the dual problem is required in the subsequent sections.

4.1. **Existence theorems.** Based on the KKM principle in the previous section, we can show under some natural assumptions that the equilibrium problem $EP(K, F)$ is solvable.

Theorem 4.1. *Suppose that $K \subset X$ is closed convex, and $F : K \times K \to \mathbb{R}$ is a bifunction satisfying the following properties:*

(A1) $F(x, x) \geq 0$ for each $x \in K$.
(A2) For every $x \in K$, the set $\{y \in K \ ; \ F(x, y) < 0\}$ is convex.
(A3) For every $y \in K$, the function $x \mapsto F(x, y)$ is usc.
(A4) There exists a compact subset $L \subset K$ containing a point $y_0 \in L$ such that $F(x, y_0) < 0$ whenever $x \in K \setminus L$.

Then, the problem $EP(K, F)$ has a solution and the solution set $\mathcal{E}(K, F)$ is closed. In addition, if the function $x \mapsto F(x, y)$ is quasi-concave for every fixed $y \in K$, then $\mathcal{E}(K, F)$ is convex.

Proof. Define a set-valued mapping $G : K \rightrightarrows K$ by

$$G(y) := \overline{\text{lev}}_{F(\cdot, y)}(0) = \{x \in K \ ; \ F(x, y) \geq 0\}, \quad \forall y \in K.$$

It is clear that $EP(K, F)$ has a solution if and only if $\bigcap_{y \in K} G(y) \neq \varnothing$. Thus, it is sufficient to show the nonemptiness of $\bigcap_{y \in K} G(y)$.

By (A1) and (A2), we may see that G has nonempty closed values at all $y \in K$. We will show now that G satisfies the inclusion (3.2). Let us suppose to the contrary that there exist $y_1, \ldots, y_m \in K$ such that $\text{co}(\{y_1, \ldots, y_m\}) \not\subset \bigcup_{j=1}^m G(y_i)$. That is, there exists a point $y^* \in \text{co}(\{y_1, \ldots, y_m\})$ such that $y^* \notin G(y_i)$ for any $i = 1, \ldots, m$. It further implies that

$$F(y^*, y_i) < 0, \quad \forall i = 1, \ldots, m.$$

Moreover, we have for all $i = 1, \ldots, m$, $y_i \in \{y \in K \ ; \ F(y^*, y) < 0\}$, which is a convex set by hypothesis. Since $y^* \in \text{co}(\{y_1, \ldots, y_m\})$ and $\text{co}(\{y_1, \ldots, y_m\})$ is the

smallest convex set containing $y_1, \ldots, y_m$, we get

$$y^* \in \mathrm{co}(\{y_1, \ldots, y_m\}) \subset \{y \in K \; ; \; F(y^*, y) < 0\},$$

which further gives $F(y^*, y^*) < 0$. This contradicts hypothesis (A1), and therefore G satisfies (3.2). On the other hand, hypothesis (A4) forces $G(y_0) \subset L$, which guarantees the compactness of $G(y_0)$. Since G satisfies every conditions of Lemma 3.3, we get $\bigcap_{x \in X} G(x) \neq \varnothing$.

Next, suppose that $(x^k) \subset \mathcal{E}(K, F)$ is convergent to $\bar{x} \in K$, then (A3) gives

$$F(\bar{x}, y) \geq \limsup_{k \longrightarrow \infty} F(x^k, y) \geq 0, \quad \forall y \in K.$$

So, we have $\bar{x} \in \mathcal{E}(K, F)$ and hence the closedness of $\mathcal{E}(K, F)$.

Now, assume that $x \mapsto F(x, y)$ is quasi-concave for every fixed $y \in K$. Let $\hat{x}, \hat{y} \in \mathcal{E}(K, F)$ and $t \in [0, 1]$. Then, we have

$$F(\gamma_{\hat{x}, \hat{y}}(t), y) \geq \min\{F(\hat{x}, y), F(\hat{y}, y)\} \geq 0, \quad \forall y \in K,$$

which implies the convexity of $\mathcal{E}(K, F)$. ∎

Next, we deduce an existence theorem for a compact convex domain.

Corollary 4.2. *Suppose that $K \subset X$ is compact convex, and $F : K \times K \to \mathbb{R}$ is a bifunction satisfying the following properties:*

(A1) $F(x, x) \geq 0$ for each $x \in K$.
(A2) For every $x \in K$, the set $\{y \in K \; ; \; F(x, y) < 0\}$ is convex.
(A3) For every $y \in K$, the function $x \mapsto F(x, y)$ is usc.

Then, the problem $EP(K, F)$ has a solution and the solution set $\mathcal{E}(K, F)$ is closed. In addition, if the function $x \mapsto F(x, y)$ is quasi-concave for every fixed $y \in K$, then $\mathcal{E}(K, F)$ is convex.

Proof. Apply Theorem 4.1 with $L = K$ in the hypothesis (A4). ∎

4.2. **Dual Equilibrium Problem.** It is essential to also introduce here the duality to the problem $EP(K, F)$ in the sense of Minty. To clarify the terminology, we shall sometimes refer to $EP(K, F)$ as the *primal problem*. The *dual equilibrium problem* to $EP(K, F)$, denoted by $EP^*(K, F)$, is given as follows:

Find a point $\bar{x} \in K$ such that $F(y, \bar{x}) \leq 0$ for every $y \in K$. $EP^*(K, F)$

The solution to $EP^*(K, F)$ will also be called the *dual solution to $EP(K, F)$*, and the set of such dual solutions is then denoted by $\mathcal{E}^*(K, F)$.

Definition 4.3. F is called *monotone* if $F(x, y) + F(y, x) \leq 0$ for all $x, y \in K$.

If F is a monotone bifunction, we immediately have the inclusion $\mathcal{E}(K, F) \subset \mathcal{E}^*(K, F)$. To obtain the converse, we need additional assumptions on the bifunction F.

Proposition 4.4. *Suppose that $F : K \times K \to \mathbb{R}$ is a bifunction.*

(i) If F is monotone, then $\mathcal{E}(K, F) \subset \mathcal{E}^*(K, F)$.
(ii) If (A1), (A3) holds, and F is convex in the second variable, then $\mathcal{E}^*(K, F) \subset \mathcal{E}(K, F)$.

In particular, if F is monotone, convex in the second variable, and satisfies (A1) and (A3), then $\mathcal{E}^*(K, F) = \mathcal{E}(K, F)$.

Proof. (i) Let $\bar{x} \in \mathcal{E}(K, F)$. The monotonicity of F then gives $F(y, \bar{x}) \leq -F(\bar{x}, y) \leq 0$, so that $\bar{x} \in \mathcal{E}^*(K, F)$.

(ii) Suppose that $\bar{x} \in \mathcal{E}^*(K, F)$. Let $y \in K$ be arbitrary. Take into account (A1), we have

$$0 \leq F(\gamma_{\bar{x},y}(t), \gamma_{\bar{x},y}(t)) \leq (1-t)F(\gamma_{\bar{x},y}(t), \bar{x}) + tF(\gamma_{\bar{x},y}(t), y) \leq F(\gamma_{\bar{x},y}(t), y),$$

for any $t \in [0, 1]$. In view of (A3), we have $0 \leq \limsup_{t \longrightarrow 0} F(\gamma_{\bar{x},y}(t), y) \leq F(\bar{x}, y)$. This gives $\bar{x} \in \mathcal{E}(K, F)$.   ∎

5. Resolvents of Bifunctions

The resolvent is a fundamental notion in regularization of certain variational problems. It is understood as the solution set of particular regularized problem from the original one. Here, we introduce the resolvent of a bifunction F in relation to the equilibrium problem $EP(K, F)$.

For simplicity, we shall adopt the following perturbation $\tilde{F}$ of a given bifunction F. That is, given a bifunction $F : K \times K \to \mathbb{R}$, $\bar{x} \in X$, we define the function $\tilde{F}_{\bar{x}} : K \times K \to \mathbb{R}$ by

$$\tilde{F}_{\bar{x}}(x, y) := F(x, y) - \langle \overrightarrow{x\bar{x}}, \overrightarrow{xy} \rangle, \quad \forall x, y \in K.$$

We are now ready to construct the resolvent of a bifunction F as the unique equilibrium of a perturbed bifuntion.

Definition 5.1 ([19]). Suppose that $K \subset X$ is closed convex, and $F : K \times K \to \mathbb{R}$. The *resolvent* of F is the mapping $J_F : X \rightrightarrows K$ defined by

$$J_F(x) := \mathcal{E}(K, \tilde{F}_x) = \{z \in K \mid F(z, y) - \langle \overrightarrow{z\bar{x}}, \overrightarrow{zy} \rangle \geq 0, \forall y \in K\}, \quad \forall x \in X.$$

The question well-definedness of J_F is important. Kimura and Kishi [19] showed a well-definedness result under the CHFP assumption. The use of CHFP assumption is to trigger the KKM principle of Niculescu and Rovenţa [28]. However, we can replace such KKM principle of Niculescu and Rovenţa [28] with our Theorem 3.3 and follows the same proof from [19] to obtain the well-definedness of J_F. Thus, we obtain Kimura and Kishi's theorem without CHFP.

Theorem 5.2. *Suppose that F has the following properties:*

(i) $F(x, x) = 0$ *for all $x \in K$.*
(ii) F *is monotone.*
(iii) *For each $x \in K$, $y \mapsto F(x, y)$ is convex and lsc.*
(iv) *For each $y \in K$, $F(x, y) \geq \limsup_{t \downarrow 0} F(\gamma_{x,z}(t), y)$ for all $x, z \in K$.*

Then $\mathrm{dom}(J_F) = X$ and J_F is single-valued.

Note that if we drop the assumption that $y \mapsto F(x, y)$ is lsc, we get the following result by imposing the flatness and local compactness.

Proposition 5.3. *Suppose that hypotheses (A1), (A3), and (A4) from Theorem 4.1 are satisfied. Suppose also that the following additional assumptions hold:*

(i) F *is monotone.*
(ii) *for any fixed $x \in K$, $y \mapsto F(x, y)$ is convex.*

If a nonempty convex subset $N \subset X$ is flat and locally compact, then the resolvent J_F is defined for all $x \in N$. That is, $N \subset \mathrm{dom}(J_F)$.

Proof. Fix $\tilde{x} \in N$, we define a bifunction $g_{\tilde{x}} : K \times K \to \mathbb{R}$ by

$$g_{\tilde{x}}(z, y) := F(z, y) - \langle \overrightarrow{z\tilde{x}}, \overrightarrow{zy} \rangle, \quad \forall z, y \in K.$$

Then $g_{\tilde{x}}(z, z) \geq 0$ for all $z \in K$, and $z \mapsto g_{\tilde{x}}(z, y)$ is usc for any fixed $y \in K$. Since K is flat, the mapping $y \mapsto \langle \overrightarrow{z\tilde{x}}, \overrightarrow{zy} \rangle$ is affine by (iii) of Lemma 2.3. Therefore, the set $\{y \in K \; ; \; g_{\tilde{x}}(z, y) < 0\}$ is convex for each $z \in K$.

Since all assumptions of Theorem 4.1 are satisfied, F has an equilibrium $\bar{x} \in K$. Now, using the monotonicity of F and (i) of Lemma 2.3, we get

$$\begin{aligned} g_{\tilde{x}}(z, \bar{x}) &= F(z, \bar{x}) - \langle \overrightarrow{z\tilde{x}}, \overrightarrow{z\bar{x}} \rangle \\ &\leq -F(\bar{x}, z) + \langle \overrightarrow{\bar{x}\tilde{x}}, \overrightarrow{\bar{x}z} \rangle - \rho^2(\bar{x}, z) \\ &\leq (\rho(\bar{x}, \tilde{x}) - \rho(\bar{x}, z))\rho(\bar{x}, z). \end{aligned}$$

Set $L := K \cap \mathrm{cl}\, A(\bar{x}; \rho(\bar{x}, \tilde{x}))$, where $\mathrm{cl}(\cdot)$ denotes the closure operator. Since K is locally compact and L is closed and bounded, we get the compactness of L. Moreover, we have $\bar{x} \in L$ and $g_{\tilde{x}}(z, \bar{x}) < 0$ for all $z \in K \setminus L$. In fact, we have just showed that $g_{\tilde{x}}$ verifies all assumptions of Theorem 4.1, and therefore $g_{\tilde{x}}$ has an equilibrium $\bar{z} \in K$. Equivalently, we have proved that $\bar{z} \in J_F(\tilde{x})$. Since $\tilde{x} \in N$ is chosen at arbitrary, the desired result is attained. ∎

The following proposition gives several valuable properties for a resolvent of a monotone bifunction.

Proposition 5.4. *Suppose that F is monotone and $\mathrm{dom}(J_F) \neq \varnothing$. Then, the following properties hold.*

    (i) J_F is single-valued.
    (ii) If $\mathrm{dom}(J_F) \supset K$, then J_F is nonexpansive restricted to K.
    (iii) If $\mathrm{dom}(J_{\mu f}) \supset K$ for any $\mu > 0$, then $Fix(J_F) = \mathcal{E}(K, F)$.

Proof. (i) Let $x \in \mathrm{dom}(J_F)$ and suppose that $z, z' \in J_F(x)$. So, we get

$$\begin{cases} F(z, z') \geq \langle \overrightarrow{z\tilde{x}}, \overrightarrow{zz'} \rangle, \\ F(z', z) \geq \langle \overrightarrow{z'x}, \overrightarrow{z'z} \rangle. \end{cases}$$

By summing up the two inequalities, applying the monotonicity of F and (i) of Lemma 2.3, we obtain

$$0 \geq F(z, z') + F(z', z) \geq \langle \overrightarrow{z\tilde{x}}, \overrightarrow{zz'} \rangle + \langle \overrightarrow{z'x}, \overrightarrow{z'z} \rangle = \rho^2(z, z').$$

This shows $z = z'$.

(ii) Let $x, y \in K$. By the definition of J_F, we have

$$\begin{cases} F(J_F(x), J_F(y)) - \langle \overrightarrow{J_F(x)x}, \overrightarrow{J_F(x)J_F(y)} \rangle \geq 0, \\ F(J_F(y), J_F(x)) - \langle \overrightarrow{J_F(y)y}, \overrightarrow{J_F(y)J_F(x)} \rangle \geq 0. \end{cases}$$

Summing the two inequalities above yields

$$\begin{aligned} 0 &\geq \langle \overrightarrow{J_F(x)x}, \overrightarrow{J_F(x)J_F(y)} \rangle + \langle \overrightarrow{J_F(y)y}, \overrightarrow{J_F(y)J_F(x)} \rangle \\ &= \left[\rho^2(J_F(x), x) + \rho^2(J_F(x), J_F(y)) - \rho^2(x, J_F(y)) \right] \\ &\quad + \left[\rho^2(J_F(y), y) + \rho^2(J_F(x), J_F(y)) - \rho^2(y, J_F(x)) \right]. \end{aligned}$$

Rearraging terms in the above inequality and apply Proposition 2.1, we get

$$\rho^2(J_F(x), J_F(y)) \leq \frac{1}{2} \left[\rho^2(x, J_F(y)) - \rho^2(y, J_F(x)) - \rho^2(x, J_F(x)) - \rho^2(y, J_F(y)) \right]$$

$$\leq \rho(x,y)\rho(J_F(x), J_F(y)).$$

This shows the nonexpansivity of J_F.

(iii) Let $x \in K$. Observe that

$$
\begin{aligned}
x \in Fix(J_F) &\iff x = J_F(x) \\
&\iff F(x,y) - \langle \overrightarrow{xx}, \overrightarrow{xy} \rangle \geq 0, \ \forall y \in K \\
&\iff F(x,y) \geq 0, \ \forall y \in K \\
&\iff x \in \mathcal{E}(K,F).
\end{aligned}
$$

$\blacksquare$

6. Bifunction Rosolvents and Other Resolvents

In this section, we consider two special cases of equilibrium problem, where the resolvent associated to a bifunction defined in Section 5 conincides with the resolvents designed for solving different variational problems. In particular, we deduce that our resolvent reduces to the Moreau-Yosida resolvent for a convex functional, and also to resolvents corresponding maximal monotone operators.

6.1. **Resolvents of convex functionals.** Always assume that $g : X \to \mathbb{R}$ is convex and lsc. The proximal of g is the operator $\text{prox}_g : X \to X$ given by

$$
\text{prox}_g(x) := \operatorname*{arg\,min}_{y \in X} \left[g(y) + \frac{1}{2}\rho^2(y,x) \right], \quad \forall x \in X.
$$

Also, recall (from [26, 10]) that a subdifferential of g is the set-valued vector field $\partial g : X \rightrightarrows X^*$ given by

$$
\partial g(x) := \{ \gamma \in T_x X \mid g(y) \geq g(x) + \langle \gamma, \overrightarrow{xy} \rangle, \ \forall y \in X \},
$$

for $x \in X$. It is shown that $\text{dom}(\partial g)$ is dense in $\text{dom}(g)$ (in this case, $\text{dom}(g) = X$) [26]. Moreover, we have the following lemma.

Lemma 6.1 ([10])**.** *Given $w, z \in X$, then $\overrightarrow{zw} \in \partial g(z)$ if and only if $z = \text{prox}_g(w)$.*

It is immediate that minimizing this functional g is a particular equilibrium problem. We make the following statement explicit only for the completeness.

Lemma 6.2. *Define $F_g : K \times K \to \mathbb{R}$ by*

$$
F_g(x,y) := g(y) - g(x), \quad \forall x, y \in K. \tag{6.1}
$$

Then, we have $\mathcal{E}(K, F_g) = \arg\min_K g$ and $J_{F_g} = \text{prox}_g$. Moreover, we have $\text{dom}(\text{prox}_g) = X$.

Proof. The fact that $\mathcal{E}(K, F_g) = \arg\min_K g$ is obvious. Now, Lemma 6.1 implies the following relations:

$$
\begin{aligned}
z = J_{F_g}(x) &\iff F_g(z,y) - \langle \overrightarrow{zx}, \overrightarrow{zy} \rangle \geq 0, \quad \forall y \in K \\
&\iff g(y) \geq g(z) + \langle \overrightarrow{zx}, \overrightarrow{zy} \rangle, \quad \forall y \in K \\
&\iff \overrightarrow{zx} \in \partial g(z) \\
&\iff z = \operatorname*{arg\,min}_{y \in X} \left[g(y) + \frac{1}{2}\rho^2(y,x) \right] = \text{prox}_g(x)
\end{aligned}
$$

The fact that $\text{dom}(\text{prox}_g) = X$ was proved in [16, 25].

$\blacksquare$

6.2. **Resolvents of monotone operators.** Following [10], the set-valued operator $A : X \rightrightarrows X^*$ is called a *vector field* if $A(x) \subset T_x X$ for every $x \in X$. Moreover, it is said to be *monotone* if $\langle x^*, \overrightarrow{xy} \rangle \leq -\langle y^*, \overrightarrow{yx} \rangle$ for every $(x, x^*), (y, y^*) \in \mathrm{grp}(A)$. It is said to be *maximally monotone* if A is monotone and $\mathrm{grp}(A)$ is not properly contained in a graph of another monotone vector field.

We shall make an observation in this section that the following *stationary problem* (or *zero point problem*) associated to a monotone set-valued vector field $A : X \rightrightarrows X^*$:

$$\text{Find a point } \bar{x} \in K \text{ such that } \mathbf{0} \in A(\bar{x}), \qquad (A^{-1}\mathbf{0})$$

can be viewed in terms of an equilibrium problem for some bifuntion F_A. The solution set is naturally expressed by $A^{-1}\mathbf{0}$.

Beforehand, let us give a restatement of a result given by [10] in the language of this paper. In fact, this states the equivalence between the singularity priblem and the Minty variational inequality of A.

Proposition 6.3. *Suppose that $A : X \rightrightarrows X^*$ is a monotone vector field with* $\mathrm{dom}(A) = X$. *Define $F_A : X \times X \to \mathbb{R}$ by*

$$F_A(x, y) := \sup_{\xi \in A(x)} \langle \xi, \overrightarrow{xy} \rangle, \quad \forall x, y \in X, \qquad (6.2)$$

Then, $A^{-1}\mathbf{0} = \mathcal{E}(X, F_A) = \mathcal{E}^*(X, F_A)$.

Proof. First, we note that the finiteness of F_A is guaranteed readily from the monotonicity of A. Moreover, the fact that $\mathcal{E}(X, F_A) \subset \mathcal{E}^*(X, F_A) = A^{-1}\mathbf{0}$ follows from the monotonicity of F_A, [10, Lemma 3.5], and [10, Lemma 3.6]. It therefore suffices to show that $A^{-1}\mathbf{0} \subset \mathcal{E}(X, F_A)$. So, let $\mathbf{0} \in A(\bar{x})$. Then

$$0 = \langle \overrightarrow{\bar{x}\bar{x}}, \overrightarrow{\bar{x}y} \rangle \leq \sup_{\bar{\xi} \in A(\bar{x})} \langle \bar{\xi}, \overrightarrow{\bar{x}y} \rangle = F_A(\bar{x}, y), \quad \forall y \in X,$$

meaning that $\bar{x} \in \mathcal{E}(X, F_A)$. ∎

Now that we have the equivalence between the three variational problems, we continue to show that their resolvents coincide, provided that A is maximally monotone. Recall from [10] that the resolvent operator of A, denoted by R_A, is defined by $R_A(x) := \{z \in X \mid \overrightarrow{zx} \in A(z)\}$, and is single-valued.

Proposition 6.4. *Suppose that $A : X \rightrightarrows X^*$ is a maximally monotone vector field with* $\mathrm{dom}(A) = X$, *and F_A is defined by* (6.2). *Then,* $J_{F_A} = R_A$.

Proof. Let $x \in X$ be given. Then, we have

$$z = R_A(x) \iff \overrightarrow{zx} \in A(z)$$
$$\implies \langle \overrightarrow{zx}, \overrightarrow{zy} \rangle \leq \sup_{\eta \in A(z)} \langle \eta, \overrightarrow{zy} \rangle, \quad \forall y \in X \qquad (6.3)$$
$$\iff z \in J_{F_A}(x).$$

The converse of (6.3) follows from the maximal monotonicity of A. ∎

7. Proximal Algorithms

It is very natural to ask about the proximal algorithm after defining a proper resolvent operator. Let us officially define the proximal algorithm for a bifunction $F : K \times K \to \mathbb{R}$ with $\mathrm{dom}(\mu F) \supset K$ for all $\mu > 0$. Let $(\lambda_k) \subset (0, \infty)$ be the step-size sequence. The *proximal algorithm* with step sizes (λ_k) started at an initial guess $x^0 \in K$ is the sequence $(x^k) \subset K$ generated by

$$x^k := J_{\lambda_k F}(x^{k-1}), \quad \forall k \in \mathbb{N}. \qquad (\text{Prox})$$

7.1. **Convergence in functional values.** Before we continue any further, let us give a small remark on the error measurement of this proximal algorithm. For instance, at each $k \in \mathbb{N}$, we can use the definition of the resolvent J_F and the Cauchy-Schwarz inequality to deduce:

$$\inf_{y \in K} F(x^{k+1}, y) \geq -\frac{1}{\lambda_k} \rho(x^{k+1}, x^k) \sup_{y \in K} \rho(x^{k+1}, y), \tag{7.1}$$

which is useful when as a stopping criterion for the convergence, especially when the rate of asymptotic regularity is known. Note that the above estimate (7.1) will later be made sharp and precise to guarantee the convergence of the proximal algorithm (see Theorem 7.3 and 7.8, for instance). If the asymptotic regularity does not hold or the rate is unknown but K is bounded, we otherwise have

$$\inf_{y \in K} F(x^{k+1}, y) \geq -\frac{1}{\lambda_k} \operatorname{diam}(K)^2, \tag{7.2}$$

which leads directly to the convergence of functional value presented in the next theorem.

Theorem 7.1. *Suppose that* $\operatorname{dom}(J_{\mu F}) \supset K$ *for all* $\mu > 0$ *and* K *is bounded closed convex. If* $\lambda_k \longrightarrow \infty$, *then the proximal algorithm generates a sequence* $(x^k) \subset K$ *such that* $\lim_k F(x^k, y) = 0$ *for any* $y \in K$.

Proof. Follows from (7.2). $\blacksquare$

7.2. **Convergence of Proximal Algorithms.** Here, we provide a convergence theorem for proximal algorithm (Prox) for step sizes (λ_k). To prove the convergence, we first show the following 'obtuse angle' property, which gives relationship between an arbitrary point $\bar{x} \in X$, the perturbed equilibrium $\tilde{x} = J_{\mu F}(\bar{x}) \in \mathcal{E}(K, \widetilde{\mu F}_{\bar{x}})$, and the exact equilibrium $x^* \in \mathcal{E}(K, F)$.

Lemma 7.2. *Assume that* F *is monotone. Let* $\bar{x} \in X$, $\mu > 0$, $\tilde{x} \in \mathcal{E}(K, \widetilde{\mu F}_{\bar{x}})$ *and* $x^* \in \mathcal{E}(K, F)$, *then* $\langle \overrightarrow{\tilde{x}\bar{x}}, \overrightarrow{\tilde{x}x^*} \rangle \leq 0$.

Proof. Since $\tilde{x} \in \mathcal{E}(K, \tilde{F}_{\bar{x}})$, we have

$$0 \leq \widetilde{\mu F}_{\bar{x}}(\tilde{x}, x^*) = \mu F(\tilde{x}, x^*) - \langle \overrightarrow{\tilde{x}\bar{x}}, \overrightarrow{\tilde{x}x^*} \rangle,$$

which implies that $\langle \overrightarrow{\tilde{x}\bar{x}}, \overrightarrow{\tilde{x}x^*} \rangle \leq \mu F(\tilde{x}, x^*)$. Now, since $x^* \in \mathcal{E}(K, F)$ and F is monotone, we get $F(y, x^*) \leq 0$, $\forall y \in K$, and particularly $F(\tilde{x}, x^*) \leq 0$. We therefore have $\langle \overrightarrow{\tilde{x}\bar{x}}, \overrightarrow{\tilde{x}x^*} \rangle \leq 0$. $\blacksquare$

Theorem 7.3. *Suppose that* F *is monotone with* $\mathcal{E}(K, F) \neq \varnothing$, Δ-*usc in the first variable, and that* $\operatorname{dom}(J_{\mu F}) \supset K$ *for all* $\mu > 0$. *Let* (λ_k) *be bounded away from* 0. *Then the proximal algorithm* (Prox) *is* Δ-*convergent to an element in* $\mathcal{E}(K, F)$ *for any initial start* $x^0 \in K$.

Proof. Let $x^0 \in K$ be an initial start and let $x^* \in \mathcal{E}(K, F)$. We can simply see that

$$\rho(x^*, x^{k+1}) = \rho(J_{\lambda_k F}(x^*), J_{\lambda_k F}(x^k)) \leq \rho(x^*, x^k),$$

which implies that (x^k) is Fejér convergent with respect to $\mathcal{E}(K, F)$. In view of Proposition 2.8, the real sequence $\rho(x^k, x^*)$ is bounded, and hence converges to some $\xi \geq 0$. Lemma 7.2 implies that

$$\rho^2(x^{k+1}, x^k) \leq \rho^2(x^k, x^*) - \rho^2(x^{k+1}, x^*).$$

Passing $k \longrightarrow \infty$, we get $\lim_k \rho(x^{k+1}, x^k) = 0$.

Now, suppose that $\hat{x} \in K$ is a Δ-accumulation point of the sequence (x^k), and also $(x^{k_j}) \subset (x^k)$ a subsequence such that $x^{k_j} \xrightarrow{\Delta} \hat{x}$. Let $y \in K$. By the construction (Prox), we have the following inequalities for any $j \in \mathbb{N}$:

$$F(x^{k_j}, y) \geq \frac{1}{\lambda_{k_j}} \langle \overrightarrow{x^{k_j} x^{k_j-1}}, \overrightarrow{x^{k_j} y} \rangle \geq -\frac{1}{\lambda_{k_j}} \rho(x^{k_j}, x^{k_j-1}) \rho(x^{k_j}, y). \qquad (7.3)$$

Recall that (x^k) is bounded (in view of Proposition 2.8) and (λ_k) is bounded away from 0. Then (7.3) gives

$$F(x^{k_j}, y) \geq -M\rho(x^{k_j}, x^{k_j-1}), \qquad (7.4)$$

for some $M > 0$. Passing $j \longrightarrow \infty$ in (7.4) and apply the Δ-upper semiconinuity of $F(\cdot, y)$, we obtain

$$F(\hat{x}, y) \geq \limsup_{j \longrightarrow \infty} F(x^{k_j}, y) \geq -M \lim_{j \longrightarrow \infty} \rho(x^{k_j}, x^{k_j-1}) = 0.$$

Since $y \in K$ is chosen arbitrarily, we conclude that $\hat{x} \in \mathcal{E}(K, F)$. Therefore, every Δ-accumulation point of (x^k) solves $EP(K, F)$. By Proposition 2.8, the sequence (x^k) is Δ-convergent to an element in $\mathcal{E}(K, F)$. ∎

Corollary 7.4 ([5, 10]). *Suppose that $g : X \to \mathbb{R}$ is convex and lsc, and that $\arg\min g \neq \varnothing$. Then, the proximal algorithm given by*

$$\begin{cases} x^0 \in X, \\ x^k := \mathrm{prox}_{\lambda_k g}(x^{k-1}), \quad \forall k \in \mathbb{N}, \end{cases}$$

is Δ-convergent to a minimizer of g, whenever $(\lambda_k) \subset (0, \infty)$ is bounded away from 0.

Proof. Consider the bifunction F_g as defined in Lemma 6.2. The monotonicity of F_g is immediate, and the fact that F_g is Δ-usc in the first variable follows by applying Lemma 2.5 to the epigraph of $F_g(x, \cdot)$ at each $x \in X$. The convergence is then a consequence of Theorem 7.3 applied to F_g, where the remaining requirements of F_g follows from Lemma 6.2. ∎

Let us look closer at the assumptions of Theorem 7.3, as well as the bifunction F_A given by (6.2) associated to some maximal monotone vector field A. One may notice that the Δ-usc of in the first variable of F_A is irrelevant. This motivates us to find the right mechanism to link between proximal algorithms for equilibrium problems and for monotone vector fields. Fortunately enough, we can deduce another convergence criteria that is capable to include [10, Theorem 5.2] as a certain special case. To do this, we need the following notion of skewed Δ-semicontinuity of F.

Definition 7.5. A bifunction $F : K \times K \to \mathbb{R}$ is said to be *skewed Δ-upper semicontinuous* (for short, *skewed Δ-usc*) if $-F(y, x^*) \geq \limsup_k F(x^k, y)$ for all $y \in K$, whenever $(x^k) \subset K$ is Δ-convergent to $x^* \in K$.

Remark 7.6. It is clear that if F is monotone and Δ-usc in the first variable, then F is skewed Δ-usc.

The next lemma shows that we can deduce that F_A is skewed Δ-usc for a monotone vector field $A : X \rightrightarrows X^*$.

Lemma 7.7. *Suppose that $A : X \rightrightarrows X^*$ is a monotone vector field with $\mathrm{dom}(A) = X$, and F_A is defined by (6.2). If X is reflexive, then F_A is skewed Δ-usc.*

Proof. Suppose that $(x^k) \subset K$ is a Δ-convergent to $x^* \in K$, and $y \in K$ be arbitrary. Then, by the monotonicity of F_A, we have

$$\limsup_{k \longrightarrow \infty} F_A(x^k, y) \leq \limsup_{k \longrightarrow \infty} [-F_A(y, x^k)]$$

$$= \limsup_{k \longrightarrow \infty} [- \sup_{\nu \in A(y)} \langle \nu, \overrightarrow{yx^k} \rangle]$$

$$\leq \limsup_{k \longrightarrow \infty} [-\langle \nu_0, \overrightarrow{yx^k} \rangle], \qquad (7.5)$$

for any $\nu_0 \in A(y)$. Suppose that $\nu_0 = \delta \overrightarrow{yu}$ for some $u \in X$ and $\delta > 0$. Then, we have

$$\langle \nu_0, \overrightarrow{yx^k} \rangle = \langle \nu_0, \overrightarrow{yx^*} \rangle + \langle \overrightarrow{x^*u}^\delta, \overrightarrow{x^*x^k} \rangle - \langle \overrightarrow{x^*y}^\delta, \overrightarrow{x^*x^k} \rangle. \qquad (7.6)$$

Combine (7.5), (7.6), and take into account the reflexivity of X and the Δ-convergence of (x^k), we obtain

$$\limsup_{k \longrightarrow \infty} F_A(x^k, y) \leq -\langle \nu_0, \overrightarrow{yx^*} \rangle.$$

Since this is true for any $\nu_0 \in A(y)$, we get

$$\limsup_{k \longrightarrow \infty} F_A(x^k, y) \leq \inf_{\nu \in T(y)} [-\langle \nu, \overrightarrow{yx^*} \rangle] = - \sup_{\nu \in T(y)} \langle \nu, \overrightarrow{yx^*} \rangle = -F_A(y, x^*). \qquad \blacksquare$$

Now that we have the motivations for skewed Δ-semicontinuity, we shall now give another convergence theorem based on this new notion of continuity. Recall that the coincidence of the primal and dual solutions of an equilibrium problem, as appeared in the following theorem, can be referenced from either Propositions 4.4 or 6.3.

Theorem 7.8. *Suppose that F is a monotone bifunction such that $\mathcal{E}(K, F) = \mathcal{E}^*(K, F) \neq \varnothing$, skewed Δ-usc, and that $\mathrm{dom}(J_{\mu F}) \supset K$ for all $\mu > 0$. Let (λ_k) be bounded away from 0. Then the proximal algorithm (Prox) is Δ-convergent to an element in $\mathcal{E}(K, F)$ for any initial start $x^0 \in K$.*

Proof. With similar proof lines to Theorem 7.3, we can also be able to obtain (7.4). Since F is skewed Δ-usc, we obtain the following inequalities by passing $j \longrightarrow \infty$:

$$-F(y, \hat{x}) \geq \limsup_{j \longrightarrow \infty} F(x^{k_j}, y) \geq 0,$$

for each $y \in K$. This means $\hat{x} \in \mathcal{E}^*(K, F) = \mathcal{E}(K, F)$, by hypothesis. Therefore, every Δ-accumulation point lies within $\mathcal{E}(K, F)$. Apply Proposition 2.8 to conclude that (x^k) is Δ-convergent to an element in $\mathcal{E}(K, F)$. $\qquad \blacksquare$

Corollary 7.9 ([10])**.** *Suppose that $A : X \rightrightarrows X^*$ is a monotone vector field with $\mathrm{dom}(A) = X$ and $\mathrm{dom}(\mu R_A) = X$ for all $\mu > 0$, and that $A^{-1}\mathbf{0} \neq \varnothing$. Assume that X is reflexive. Then, the proximal algorithm defined by*

$$\begin{cases} x^0 \in X, \\ x^k := R_{\lambda_k A}(x^{k-1}), \quad \forall k \in \mathbb{N}, \end{cases}$$

is Δ-convergent to an element in $A^{-1}\mathbf{0}$, whenever $(\lambda_k) \subset (0, \infty)$ is bounded away from 0.

Proof. It was proved in [10] that the surjectivity condition of A implies the maximal monotonicity and also gives $\mathrm{dom}(J_{\mu F_A}) = X$ (for all $\mu > 0$) in light of Proposition 6.4, where F_A defined in (6.2). The monotonicity of F_A follows from the monotonicity of A, while the skewed Δ-semicontinuity of F_A is proved in Lemma 7.7. By the

Proposition 6.3, it is feasible to apply Theorem 7.8 to F_A and obtain the desired convergence result. ∎

Remark 7.10. Amounts to (7.1) and the Fejér convergence of the proximal algorithm presented in the proofs of Theorems 7.3 and 7.8, it can be seen that we can make the convergence of functional values $\inf_{y \in K} F(x^k, y) \longrightarrow 0$ arbitrarily fast by the choosing appropriate step sizes. However, this does not ensure the speed of convergence of the sequence (x^k) even if the convergence is strong (i.e., in the metric topology). In particular, the speed enhancement of proximal algorithm for singularity problem using metric regularity conditions was developed in [10].

Conclusion and Remarks

We have provided a complete treatment of equilibrium problem situated in Hadamard spaces, from existence to approximation algorithm. We were also able to prove the KKM principle without additional assumptions, as opposed to earlier results in the literature. The existence criteria for an equilibrium is deduced under standard assumptions, which is very natural in this subject area.

In the approximation, we investigated a priori on the resolvent operator for a given bifunction. Main results here are that the resolvent operator is single-valued and firmly nonexpansive. We then define the proximal algorithm by iterating the resolvents of different bifurcating parameters. Several convergence criteria of the algorithm were proposed, including the one that involves skewed continuity, where we introduced here for the first time. Note again that the corresponding functional values converging to 0 can be achieved arbitrarily fast, but this does not imply the same for proximal algorithm itself.

Let us conclude this paper with some open questions whose answers might largely improve the applicability of the results in this present paper.

Question 7.11. Whether or not we can improve the following condition: $\mathrm{dom}(J_{\mu F}) \supset K$ for all $\mu > 0$, in order to obtain similar results regarding resolvent operators and proximal algorithms?

Question 7.12. Is it possible to drop the surjectivity condition above and still obtain the nonexpansivity of J_F (see (ii) in Proposition 5.4)?

References

[1] B. Ahmadi Kakavandi. Weak topologies in complete CAT(0) metric spaces. *Proc. Amer. Math. Soc.*, 141(3):1029–1039, 2013.

[2] B. Ahmadi Kakavandi and M. Amini. Duality and subdifferential for convex functions on complete CAT(0) metric spaces. *Nonlinear Anal.*, 73(10):3450–3455, 2010.

[3] W. Anakkamatee and S. Dhompongsa. Rodé's theorem on common fixed points of semigroup of nonexpansive mappings in CAT(0) spaces. *Fixed Point Theory Appl.*, pages 2011:34, 14, 2011.

[4] M. Baˇcák. *Convex analysis and optimization in Hadamard spaces*, volume 22 of *De Gruyter Series in Nonlinear Analysis and Applications*. De Gruyter, Berlin, 2014.

[5] M. Bačák. The proximal point algorithm in metric spaces. *Israel J. Math.*, 194(2):689–701, 2013.

[6] I. D. Berg and I. G. Nikolaev. On a distance between directions in an Aleksandrov space of curvature $\leq K$. *Michigan Math. J.*, 45(2):257–289, 1998.

[7] I. D. Berg and I. G. Nikolaev. Quasilinearization and curvature of Aleksandrov spaces. *Geom. Dedicata*, 133:195–218, 2008.

[8] E. Blum and W. Oettli. From optimization and variational inequalities to equilibrium problems. *Math. Student*, 63(1-4):123–145, 1994.

[9] P. Chaipunya and P. Kumam. Nonself KKM maps and corresponding theorems in Hadamard manifolds. *Appl. Gen. Topol.*, 16(1):37–44, 2015.

[10] P. Chaipunya and P. Kumam. On the proximal point method in Hadamard spaces. *Optimization*, 66(10):1647–1665, 2017.

[11] V. Colao, G. López, G. Marino, and V. Martín-Márquez. Equilibrium problems in Hadamard manifolds. *J. Math. Anal. Appl.*, 388(1):61–77, 2012.

[12] S. Dhompongsa, W. A. Kirk, and B. Sims. Fixed points of uniformly Lipschitzian mappings. *Nonlinear Anal.*, 65(4):762–772, 2006.

[13] S. Dhompongsa and B. Panyanak. On Δ-convergence theorems in CAT(0) spaces. *Comput. Math. Appl.*, 56(10):2572–2579, 2008.

[14] K. Fan. A generalization of Tychonoff's fixed point theorem. *Math. Ann.*, 142:305–310, 1960/1961.

[15] J. Jost. Equilibrium maps between metric spaces. *Calc. Var. Partial Differential Equations*, 2(2):173–204, 1994.

[16] J. Jost. Convex functionals and generalized harmonic maps into spaces of nonpositive curvature. *Comment. Math. Helv.*, 70(4):659–673, 1995.

[17] H. Khatibzadeh and S. Ranjbar. A variational inequality in complete CAT(0) spaces. *J. Fixed Point Theory Appl.*, 17(3):557–574, 2015.

[18] H. Khatibzadeh and S. Ranjbar. Monotone operators and the proximal point algorithm in complete CAT(0) metric spaces. *J. Aust. Math. Soc.*, 103(1):70–90, 2017.

[19] Y. Kimura and Y. Kishi. Equilibrium problems and their resolvents in Hadamard spaces. *Proceedings of the Fifth Asian Conference on Nonlinear Analysis and Optimization (Niigata, Japan)*, 2016, to appear.

[20] Y. Kimura, S. Saejung, and P. Yotkaew. The Mann algorithm in a complete geodesic space with curvature bounded above. *Fixed Point Theory Appl.*, pages 2013:336, 13, 2013.

[21] W. A. Kirk and B. Panyanak. A concept of convergence in geodesic spaces. *Nonlinear Anal.*, 68(12):3689–3696, 2008.

[22] B. Knaster, C. Kuratowski, and S. Mazurkiewicz. Ein Beweis des Fixpunktsatzes für *n*-dimensionale Simplexe. *Fundam. Math.*, 14:132–137, 1929.

[23] T. C. Lim. Remarks on some fixed point theorems. *Proc. Amer. Math. Soc.*, 60:179–182 (1977), 1976.

[24] B. Martinet. Régularisation d'inéquations variationnelles par approximations successives. *Rev. Française Informat. Recherche Opérationnelle*, 4(Sér. R-3):154–158, 1970.

[25] U. F. Mayer. Gradient flows on nonpositively curved metric spaces and harmonic maps. *Comm. Anal. Geom.*, 6(2):199–253, 1998.

[26] M. Movahedi, D. Behmardi, and S. Hosseini. On the density theorem for the subdifferential of convex functions on Hadamard spaces. *Pacific J. Math.*, 276(2):437–447, 2015.

[27] S. Z. Németh. Variational inequalities on Hadamard manifolds. *Nonlinear Anal.*, 52(5):1491–1498, 2003.

[28] C. P. Niculescu and I. Rovenţa. Fan's inequality in geodesic spaces. *Appl. Math. Lett.*, 22(10):1529–1533, 2009.

[29] C. P. Niculescu and I. Rovenţa. Schauder fixed point theorem in spaces with global nonpositive curvature. *Fixed Point Theory Appl.*, pages Art. ID 906727, 8, 2009.

[30] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optimization*, 14(5):877–898, 1976.